

Class 9: Experiments and Observational Studies (Text: Section 3.1)

A statistical study can tell us

- What is going on (the characteristics of a population)
- Effect of an intervention (taking a drug, using a new fertilizer, implementing a policy to create jobs)

Studies may be of two types:

- **Observational study:** Characteristics of a population are observed, but the population is not disturbed.
- **Experiment:** Treatment is given to a individuals and their responses recorded.

In an experiment, the treatment is generally applied to **randomly selected** subset of the population, and the subset compared to a randomly selected **control group**. This is called a **RCT (randomized controlled trial)**. In medicine, RCTs are the gold standard for clinical trials of new drugs.

Ex. To investigate the safety of the anesthetics used in surgery, records from 850,000 operations in 34 major hospitals were examined. The death rates for four common anesthetics are shown in the table.¹

Anesthetic	A	B	C	D
Death rate	1.7%	1.7%	3.4%	1.7%

The anesthetic used is the **explanatory** variable and the death rate is the **response** variable. Was this an observational study or an experiment? What can you conclude from it?

Ex: If you have to go to hospital, should you choose the one with the lowest mortality rates? (These rates are published.)

Ex. The Physicians' Health Study (1980-1995) was a longitudinal study of the effect of aspirin and beta carotene on heart attacks.² In the study, 21,996 male physicians were randomly divided into four equal groups and given pills or placebos of aspirin and beta carotene (real versions of both, just one, or neither). The result was that 239 of the aspirin placebo group and 139 of the aspirin group had heart attacks.³ Was this an observational study or an experiment? What can you conclude from it?

Ex. A Sept 2006 article⁴, "Vitamin D halves the risk of pancreatic cancer," reported on the work scientists at Northwestern and Harvard who examined data from two long-term health studies of 46,771 men and 75,427 women. They found that people who took vitamin D had a 43% lower risk of pancreatic cancer. Comment.

1 From *Introduction to the Practice of Statistics*, by D. Moore and G. McCabe 4th edn.

2 Physicians' Health Study II was launched in 1997 to study the effects of Vitamin E, C and multivitamins.

3 Moore and McCabe, *Introduction to the Practice of Statistics*, 4th edn. "Aspirin in the primary and secondary prevention of vascular disease" published in *The Lancet* on May 30, 2009 provided a more nuanced view of the effectiveness of aspirin.

4 By Halcyon Skinner et al in MedlinePlus at NIH http://www.hpnonline.com/dailyupdates/September_06.html

Why Are Experiments/RCTs Useful? Causation

Even if the sample is well chosen, an observational study cannot easily establish causation. In an observational study there may be **hidden** variables, which explain the effect observed, but are not among the variables studied. We say that two variables are **confounded** if their effects cannot be distinguished from each other. If well designed, experiments *can* provide evidence of causation.

Ex: Children with larger shoe sizes read better.

Ex: Countries with more TV sets per person have longer life expectancy.

Ex. Oklahoma and West Virginia adopted state-funded programs promoting marriage. These are based on studies done in Oklahoma in 1998, from which⁵

“state economists concluded that being single and being poor were interrelated”.

The programs were aimed at poor single mothers:

“The hope is that marriage will provide these women, and their children, with a route out of poverty and alleviate the burden they place on the state. The thesis is simple: If failure to marry—or divorce—means poverty, marriage must enhance wealth.”

Are the 1998 studies observational studies or experiments? Can you conclude that being single causes poverty?

Ex. On Sept 3, 2002, a New York Times article, “Sorting through the Confusion over Estrogen”, reported

“.....in the decades since its initial approval by the Food and Drug Administration in 1942, estrogen had acquired a reputation as an antidote to many of the illnesses and afflictions of aging. Scores of observational and case studies supported this view.....Major new findings from such clinical trials have seriously challenged estrogen’s image as a preventive of chronic disease....”

What might have caused the results to be so different?

Multiple Regression is a primary tool for establishing causation. Also **natural experiments**, when available, as in the next example.

Ex: The Harlem Children’s Zone is an NGO (non-governmental organization) providing social services for poor children in New York City. It runs a charter school, Promise Academy, which accepts students by lottery. A study compared students who went to the school to comparable students who went to other schools.⁶ Students in the school gained 1.3 or 1.4 standard deviations (usual gains are about 0.2 SD.)

What kind of study is this?

⁵ From “Wedded to the Value of Marriage”, BBC News, January 15, 2004.

⁶ “The Harlem Miracle”, May 7, 2009, <http://www.nytimes.com/2009/05/08/opinion/08brooks.html>

Experimental Methods: Good Design

Experiments should be designed to avoid confounding and Hawthorne and experimenter effects.

Confounding is made less likely by having a

- **Control group** and
- **Random assignment** to groups.

Ex: “Learning in an online format versus and in-class format: an experimental study”⁷ reported that in Nova Southeastern University, FL, online students did as well as those who learned in class. What might you want to ask about the experiment before investing in online learning?

Ex: A 2005 experiment⁸ claimed that a drink containing pig whipworm eggs reduced the symptoms of Crohn’s disease (a disease of the bowel). What might you ask?

Experimenter and Hawthorne effects

The experimenter effect occurs when the experimenter influences the result of the experiment (by accident). The Hawthorne effect occurs when the subjects respond differently just because they are part of an experiment. Both can be eliminated by using a **placebo**.

Ex: Crohn’s disease experiment:

To make sure the experimenters’ or subjects’ behavior or beliefs do not unconsciously affect the outcome, use **double blinding** (neither experimenter nor subject knows who is in the control group) or **single blinding** (one knows; the other does not).

Ex: The Physicians’ Health Study experiment

Ex: In 2007, doctors in Germany showed that people who ate a small amount of dark chocolate each day had lower blood pressure than those who ate white chocolate.⁹ What kind of blinding was possible?

Ex: Are some “natural remedies” placebos?

7 Allan Schulman and Randi Sims, *T.H.E Journal* June 1999.

8 R. Summers et al, “Trichuris suis therapy in Crohn’s disease” in *Gut* v. 54, 2005.

9 www.medicinenet.com/script/main/art.asp?articlekey=82299 “Dark Chocolate May Help Blood Pressure” July 3, 2007

The MMR Vaccine Controversy

Children in several countries are routinely vaccinated against measles, mumps, and rubella by the MMR vaccine. As a result, measles was close to eradication. However, a 1998 paper¹⁰ linking the MMR vaccine to autism caused a considerable reduction in vaccination levels, triggering new outbreaks of measles in the US, Europe,¹¹ Israel, Japan. The 1998 paper suggested the MMR vaccine led to bowel infections and to autism. What statistical questions would help you evaluate the results of the paper?¹²

What was the research question?

Was this an observational study or an experiment?

- How big was the sample?
How was the sample chosen?

Was there a control group?

Was the study blinded?

- How long was study?
How was the study funded?

What has later research shown?¹³

What legal action has been taken?¹⁴

- British Medical Journal, 5 Jan 2011 article:¹⁵

10 "Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children" by A J Wakefield, S H Murch, A Anthony, J Linnell, D M Casson, M Malik, M Berelowitz, A P Dhillon, M A Thomson, P Harvey, A Valentine, S E Davies, J A Walker-Smith in *The Lancet*, Vol 351, No. 9103, 28 Feb 1998

11 <http://news.bbc.co.uk/2/hi/health/7813149.stm> and <http://news.bbc.co.uk/2/hi/health/8293149.stm>

12 <http://briandeer.com/mmr/lancet-greenhalgh.htm> and www.timesonline.co.uk/tol/life_and_style/health/article5683671.ece

13 www.nytimes.com/2008/08/24/opinion/24sun2.html and www.nytimes.com/2008/09/09/opinion/09tue3.html

http://www.csicop.org/si/show/autism-vaccine_link_researcher_andrew_wakefield_accused_of_faking_his_data/

14 <http://jolt.unc.edu/blog/2009/02/18/no-link-between-mmr-vaccine-and-autism> and <http://news.bbc.co.uk/2/hi/health/8695267.stm>

15 <http://www.bmj.com/content/342/bmj.c7452.full>

Ex: In March 2009 in Minneapolis, 25% of the children in preschool classes for children with severe autism were Somali, while 6% of the school enrollment was Somali.¹⁶ What can you conclude?

Ex: On September 22 2009, a study¹⁷ compared adult and child autism rates in UK. What might you expect to see? What light could this shed on the safety of the MMR vaccine?

What is an Institutional Review Board? (IRB) Why were they created?

Tuskegee Study

“... study that was conducted between 1932 and 1972 in Tuskegee, Alabama. That study included nearly 400 poor African-American men with preexisting syphilis whose disease was allowed to progress without treatment. Researchers did not infect the subjects, but they did not tell them they had the disease either.”¹⁸ The men thought they were receiving free health care from the US government, but they were not given treatment even after penicillin was discovered to be a cure for syphilis in the 1940s.” As a result, some died of syphilis, wives got syphilis and children were born with congenital syphilis. Study was brought to light and stopped in 1972.

Guatemala Study

“...a 1946-1948 research study in which people in Guatemala were intentionally infected with sexually transmitted diseases..... the U.S. Public Health Service Sexually Transmitted Disease Inoculation Study of 1946-1948, aimed at determining the effectiveness of penicillin in treating or preventing syphilis after subjects were exposed to the disease. Gonorrhea and chancres were also studied. Penicillin was a relatively new drug at the time. The tests were carried out on female commercial sex workers, prisoners in the national penitentiary, patients in the national mental hospital and soldiers. According to the study, more than 1,600 people were infected: 696 with syphilis, 772 with gonorrhea and 142 with chancres.”¹⁹

An IRB must give approval to all research involving human subjects in its institution. The IRB considers:

Can an observational study require IRB supervision?

¹⁶ http://www.nytimes.com/2009/03/17/health/17auti.html?pagewanted=1&_r=2&hp

¹⁷ “Autism Spectrum Disorders in adults living in households throughout England - report from the Adult Psychiatric Morbidity Survey 2007” Sept 22, 2009, The NHS Information Centre

¹⁸ <http://edition.cnn.com/2010/WORLD/americas/10/01/us.guatemala.apology/index.html?iref=obinsite>

¹⁹ <http://edition.cnn.com/2010/WORLD/americas/10/01/us.guatemala.apology/index.html?iref=obinsite>